

Durham Research Online

Deposited in DRO:

05 February 2016

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Abdelsalam, O. and Duygun, D. and Matallín-Sáez, J.C. and Tortosa-Ausina, E. (2014) 'Do ethics imply persistence? The case of Islamic and socially responsible funds.', *Journal of banking finance.*, 40 . pp. 182-194.

Further information on publisher's website:

<http://dx.doi.org/10.1016/j.jbankfin.2013.11.027>

Publisher's copyright statement:

© 2013 This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Do ethics imply persistence? The case of Islamic and socially responsible funds^{*}

Omneya Abdelsalam
Aston University

Meryem Duygun
University of Leicester

Juan Carlos Matallín
Universitat Jaume I

Emili Tortosa-Ausina
Universitat Jaume I and Ivie

January 16, 2014

Abstract

We analyse the performance persistence of Islamic and Socially Responsible Investment (SRI) mutual funds. We adopt a multi-stage strategy in which, in the first stage, partial frontiers' approaches are considered to measure the performance of the different funds in the sample. In the second stage, the results yielded by the partial frontiers are plugged into different investment strategies based on a recursive estimation methodology whose persistence performance is evaluated in the third stage of the analysis. Results indicate that, for both types of funds, performance persistence actually exists, but only for the worst and, most notably, best funds. This result is robust not only across methods (and different choices of tuning parameters within each method) but also across both SRI and Islamic funds—although in the case of the latter persistence was stronger for the best funds. The persistence of SRI and Islamic funds represents an important result for investors and the market, since it provides information on both which funds to invest in and which funds to avoid. Last but not least, the use of the aforementioned techniques in the context of mutual funds could also be of interest for the non-conclusive literature.

Keywords: Islamic funds, Socially Responsible Investment funds, performance, persistence

JEL Classification: F15, F21, F36, Z13

Communications to: Emili Tortosa-Ausina, Departament d'Economia, Universitat Jaume I, Campus del Riu Sec, 12071 Castelló de la Plana, Spain. Tel.: +34 964387168, fax: +34 964728591, e-mail: tortosa@uji.es

^{*}All authors are grateful to an anonymous referee for helpful comments, as well as to El Shaarani Centre for Islamic Business and Finance for financially supporting this project. Juan Carlos Matallín and Emili Tortosa-Ausina acknowledge the financial support provided by research projects ECO2011-27227/ECON (Ministerio de Ciencia e Innovación), P1-1B2009-54 (Fundació Caixa Castelló-Bancaixa and Universitat Jaume I) and PROM-ETEO/2009/066 (Generalitat Valenciana). Emili Tortosa-Ausina is also thankful to the El Shaarani Centre for Islamic Business and Finance and the Aston Business School for their hospitality.

1. Introduction

Over the last twenty years a relatively high number of studies have dealt with analyzing persistence in mutual funds' performance. The importance of the topic is related to how relevant it might be to include past performance as a key criterion for selecting mutual funds. Given the importance of this question, i.e. whether good past performance is followed by good subsequent performance (and likewise for bad performance), for both academics and practitioners, the related literature has grown remarkably over the last two decades.

Until the mid 1990s, several contributions provided competent efforts to document the issue, although, as indicated by Carhart (1997), explanations were generally not entirely satisfactory. Some of these studies would include, for instance, Hendricks et al. (1993), Goetzmann and Ibbotson (1994) and Brown and Goetzmann (1995), who attributed the persistence found over short-term horizons (up to three years) to common investment strategies (i.e. "hot hands"), whereas Grinblatt and Titman (1992), Elton et al. (1993) and Elton et al. (1996) found it over longer time horizons (five to ten years), attributing it to managerial stock-picking skills. Although most of the influential initial studies were published in the early and mid 1990s, a few contributions, such as Jensen (1969), had previously contributed to the topic, finding no evidence of persistence.

The study by Carhart (1997) has been perhaps the most influential and successful in finding some convincing explanations related to the existence of persistence found in the literature. According to this author, persistence could be caused by managers' costs and the momentum effect, rather than managers' skills. With the exception of the worst performing funds, Carhart's study revealed no significant evidence of persistence in his analysis of investment strategies based on past performance.

However, the literature has not stood still and the most recent contributions have not reached a consensus as to the lack or existence of persistence, as well as its potential explanations. Some more recent studies such as Quigley and Siquefield (2000), ? and Massa and Patgiri (2009) found similar conclusions to those by Carhart (1997) in terms of no significant evidence of persistence in the analysis of investment strategies based on past performance. In contrast, Lynch and Musto (2003), Cohen et al. (2005), or Kosowski et al. (2006) found persistence among winning funds, but not among losers. Chen et al. (2000), Cohen et al. (2005) or Cremers and Petajisto (2009) have also found evidence of persistence and others report a list of causes as well. These studies would include, for instance, Gottesman and Morey (2007), who consider persistence to be caused mainly by the expense ratio, Fama and French (2010), who identify costs as the main source of persistence and Bollen and Busse (2004), who found persistence beyond expenses and momentum. Wermers (2003), by examining managers' momentum, finds evidence of persistence in superior growth funds, whereas in the case of Kosowski et al. (2006), using net returns after trading costs and fees, they find evidence of persistence in growth-oriented funds, but find no evidence for the managerial skills of income-oriented funds.

A good review of this important literature is provided by Droms (2006), who summarizes conclusions one might draw after reviewing the relevant literature, among which we might highlight, above all, that past performance counts, i.e. persistence exists, at least in the short run. However, according to Droms (2006), the literature has also found that: (i) poor past performance counts more than good past performance; (ii) within-category performance is more persistent than performance relative to overall market; (iii) short term performance persistence is much stronger than long-term persistence; (iv) international equity funds show very strong short-term performance persistence but no longer-term persistence; (v) findings of persistence may be sensitive to the period tested; and that (vi) Morningstar within category ratings provide useful information for selecting mutual funds based on past performance (Droms, 2006).

In the last few decades, a particular branch of the literature on mutual funds' performance has been interested in analyzing ethically oriented funds, a number of which have become very

popular. Investors in ethical funds (such as Socially Responsible Funds and Islamic Funds) apply both financial and social criteria when evaluating their investments in order to ensure that the securities selected are consistent with their personal value system and beliefs (Sauer, 1997). As different ethical funds exhibit distinct investment styles (Bauer et al., 2005) and apply varying screening criteria (Derigs and Marzban, 2008), the diverse constraints applied are likely to influence the performance and performance persistence of different types of ethical funds.

In the specific case of the literature analyzing ethical funds and SRI performance *persistence*, the number of available studies is low (compared with the literature focusing on conventional funds) and results provide mixed empirical evidence. While Sauer (1997) finds that well diversified benchmark portfolios suggest that application of social responsibility screens does not necessarily have an adverse impact on investment performance, Gregory and Whittaker (2007), covering UK SRI funds, provide empirical evidence that supports persistence in performance, particularly over longer time horizons. However, they find that performance appears to be time and model-varying. They also find some evidence that for domestic funds, past “winning” SRI responsibility screens do not necessarily have an adverse impact on investment performance and flow. A number of other studies found performance persistence for ethical or SRI funds (Gregory and Whittaker, 2007; Renneboog et al., 2011). Bauer et al. (2005) find that ethical mutual funds are typically less exposed to market return variability compared to conventional funds, which could lead to persistence. Renneboog et al. (2011) find that ethical money is less sensitive to past negative funds, outperforming “losing” SRI funds to a greater extent than their control portfolio counterparts.

For Islamic funds, a handful of studies compared the performance of these funds with their conventional counterparts (Ahmad and Ibrahim, 2002; Girard and Hassan, 2008; Hashim, 2008; Albaity and Ahmad, 2008; Dharani and Natarajan, 2011; Mansor and Bhatti, 2011; BinMahfouz and Hassan, 2012). Most of the studies found that there is no significant difference in the performance of restricted Islamic funds from their conventional counterparts except Hussein and Omran’s (2005) study, in which Islamic indices performed better than their conventional equivalents. When it comes to explicitly analyzing the *persistence* in Islamic mutual fund performance the literature is yet to be written. Therefore, our study combines two relevant stems of research in mutual fund performance evaluation, in which the evidence is either scarce or entirely yet to come. Namely, we conduct a performance persistence analysis for two particular types of ethical funds—SRI funds and Islamic funds. However, there are added reasons justifying undertaking this analysis. As indicated in the paragraph above, Islamic mutual funds have very specific features such as investing in *Shari’ah* compliant assets, as well as other relevant characteristics, which bear some similarities with others attributable to SRI funds. It could therefore be hypothesized that investing in some specific types of assets could have a payoff in terms of higher persistence, especially for the best performers.

These two types of funds use restricted universes due to strict screening criteria. As a result, their constituents probably enjoy more stability compared to their conventional counterparts due to the stability in the filtering criteria (Askari et al., 2010). This stability in the constituents might lead to persistence in the performance of these funds. In addition, in the case of Islamic funds, they apply further financial ratio filters on the equity selected, such as leverage and percentage of interest paid or received, which restrict their universe further. Moreover, Bauer et al. (2005) find that ethical mutual funds are typically less exposed to market return variability compared to conventional funds which could lead to persistence.

Our empirical strategy will be based on a *recursive* estimation methodology, similar to that initially proposed by Carhart (1997)—although several refinements were proposed afterwards by Bollen and Busse (2004), Kosowski et al. (2006), Busse et al. (2010) and Fama and French (2010), among others. The main purpose of this method is to ascertain whether performance

persistence exists based on the evaluation of portfolios which are built following a past performance investment strategy. Should persistence exist, past performance investment strategies would succeed—i.e. the best funds remain so the year after and vice versa.

In a previous stage, we explicitly measured mutual fund performance, for which the literature has grown remarkably over the last two decades. However, in more recent times, the approaches based on frontier methodologies, such as Data Envelopment Analysis (DEA), have flourished. Both DEA and its non-convex variant, FDH (Free Disposal Hull) share the attractive feature of easily accommodating several dimensions of performance (inputs and outputs), instead of confining the analysis to a few attributes (e.g. mean and variance), although they are not free from criticisms. In particular, they may be severely affected by outliers and the so-called “curse of dimensionality”, but more recent proposals in the field such as the order- m (Cazals et al., 2002) and order- α (Aragon et al., 2005) (which we will follow) are much less affected by these and related issues.

The combination of our recursive estimation methodology for measuring performance persistence with the robust nonparametric estimators proposed by Cazals et al. (2002) is particularly interesting, since it provides information not only as to whether performance persists for different parts of the funds’ performance distribution (not only the upper and lower tails), but also whether this persistence is actually significant or not. As we shall see, although persistence is actually found for several deciles of the distributions, only for the best and worst funds (especially in the case of SRI funds) is it found to be significant. This result, which was partly found in previous studies such as Goetzmann and Ibbotson (1994), also held for the different specifications of tuning parameters allowed by the nonparametric robust techniques used in the first stage.

The article is structured as follows: after this introduction, section 2 briefly reviews the literature in the field of SRI and Islamic mutual fund performance. Section 3 presents the details of the models and methods that we use to both evaluate performance and its persistence, whereas section 4 and 5 describe the data and the main results, respectively. Finally, section 6 presents some concluding remarks.

2. SRI and Islamic funds

Over the past two decades, some types of mutual funds have grown exponentially. This includes socially responsible investment (SRI) and faith-based mutual funds which have paralleled the growth in the business ethics literature as a result of recent reports of environmental and accounting scandals. These funds use techniques that combine investors’ financial objectives with their commitments to social and/or religious concerns (Hiagh and Hazelton, 2004). These types of funds apply ethical restrictions on investments which can adversely affect portfolio performance and performance persistence (Bauer et al., 2007).

The earliest SRI screening was applied by religious groups such as the Lutheran Brotherhood and the Quakers which excluded “sin industries” such as tobacco and alcohol (Schepers, 2003). The PAX world fund was launched in 1971 as the first SRI mutual fund which avoided investments in military-related stocks (Fowler and Hope, 2007). A number of other funds with social and peace-driven goals have been initiated leading professionally managed SRI assets in the U.S. to reach \$3.07 trillion at the start of 2010 (Social Investment Forum, 2010) and EUR 129.49 billion in Europe as of June 2012.

According to Ernst & Young (2011), Islamic mutual funds to the tune of \$1,033 billion are managed, making this type of finance the fastest growing of its kind within the Islamic financial industry. Dating back to 1994, Islamic equity investment is a new phenomenon. In 1994 Muslim investors, under a new ruling¹ were permitted to trade in international stocks, within prescribed

¹The decree issued by International Fiqh Academy relaxed the *Shari’ah* constraints on interest-based activities

guidelines (Hayat and Kraeussl, 2011). From then on, Islamic investment has increased exponentially, with many global players subscribing to the Islamic investment market. By the end of 2010, FTSE, Dow Jones, S&P and MSCI had been offering Islamic equity indices in the hundreds, leading to over 800 managed Islamic mutual funds according to Ernst & Young (2011).

Four pillars represent the defining prohibitions of Islamic finance, namely *Gharar* (excessive uncertainty); *Riba* (usury);² *Maysir* (speculation); and investing in prohibited activities. The fifth pillar is the encouragement of risk and return sharing (Shanmugam and Zahari, 2009; Hayat and Kraeussl, 2011). By Islamic standards, charging a fixed rate of interest on an investment loan is deemed unfair and discriminatory, as the borrower, or entrepreneur, shoulders the full risk and the lender, whether or not the venture succeeds, gains the set income. Conversely, when there is a very high profit, the lender will receive a relatively low portion of the profit, while the borrower will receive the *lion's* share implying an unequal sharing of both risk and profit (Novethic, 2009). A panel of *Shari'ah* experts, the *Shari'ah* Supervisory Board or SSB, keeps a close check on the compliance of funds and companies with the strictures of *Shari'ah*.

As explained by Forte and Miglietta (2007), both types of funds have social, ethical, and financial objectives, and both use negative screening (filtering) criteria in the selection of stocks for their portfolios to reflect these objectives. For Islamic funds, the source of guidance is *Shari'ah*. However, due to the lack of a global *Shari'ah* supervisory body, different funds appoint their own *Shari'ah* Supervisory board (SSB) and apply differing interpretations of *Shari'ah* screening criteria.³ For SRI, they were historically derived from religious factions that evaded investment in stocks considered "Sin" reserves. Nonetheless, the idea of SRI was gradually developed and was additionally advanced and expanded by ecological policies, human rights filters and anti-war projects. There are no comprehensively recognised definitions for Sustainable and Responsible Investment (SRI) and Environmental, Social and Governance (ESG) principles.

In terms of investment activities and instruments, Islamic funds exclude investments in instruments with fixed income, such as certificates of deposit (CDs), corporate bonds, preferred stocks and some derivatives (e.g., options). While equity mutual funds represent the largest slice of Islamic funds, SRI mutual funds can freely choose between equity-bearing investments and debt-bearing investments, as long as the chosen stocks adhere to sustainable and responsible investment strategies and governance principles.

In addition, Islamic funds apply further financial ratio filters on the equity selected, such as leverage and percentage of interest paid or received.⁴ There is no positive screening practiced by Islamic funds. However, SRI funds apply different strategies for positive filters, such as best in class environmental filters, human rights and transparent corporate practices.

According to portfolio theory, investors choose portfolios that maximize returns while minimizing risks, thereby receiving the best possible return on their investment. Constraining funds by ethical and religious filters is likely to influence the returns and risks of restricted portfolios (Binmahfouz, 2012). Ongoing research is being conducted and analysis made of the various investment traits of restricted Islamic funds, thereby delineating their difference in performance

for companies whose main industry sector is permissible (*Halal*) (Hayat and Kraeussl, 2011).

²*Riba* is prohibited in "all monotheist religions" (Hossain, 2009).

³The research by Derigs and Marzban (2008) has illustrated that there is a lack of a commonly recognised comprehension of how to convert the prescribed *Shari'ah* regulations into a structure of investment guidelines that can be monitored. The research has additionally illustrated that implementing the alternative *Shari'ah* screens, employed by the more renowned Islamic indexes and funds that are *Shari'ah*-compliant to a general standard asset universe (S&P500 index), has caused considerable variations in the size as well as the constituents of the resultant portfolios of halal assets. For the purpose of dealing with the discrepancy it is thus essential for a *Shari'ah* authority of global standing to be established, so as to address the variation of the *Shari'ah* screening guidelines.

⁴A purification process must be carried out to eliminate or clean the portfolio of interest income, or other impermissible revenue sources. Impermissible income has to be donated to charities and NGOs. However, for SRI there is no purification process necessary.

from unrestricted equivalents.

3. Methods

As indicated by Kerstens et al. (2011b), the theoretical foundations of modern portfolio theory—which combine the expected return and variance data to assess performance—have been subject to criticisms, from both theoretical and applied points of view including, for instance, strong assumptions on probability distributions and Von Neumann-Morgenstern utility functions. Since the foundations of traditional performance measurement were established, a vast literature has emerged on portfolio performance measurement.

This literature has moved gradually from total-risk foundations to performance indexes (where returns in excess of the risk-free rate are compared to some risk measure (Kerstens et al., 2011b)). However, in the particular (and relevant) case of mutual funds, there is still no universally accepted assessment approach to measure their performance. In a relatively recent study, Eling and Schuhmacher (2007) reviewed not only the most widely-used measures, namely Sharpe’s, Treynor’s and Jensen’s, along with multifactor alpha measures, but also other measures which are less popular but have had varying degrees of acceptance in recent times.⁵

In relatively recent times, partly in view of the criticisms of traditional performance measures, some scholars and practitioners have been applying the so-called frontier estimation methodologies from production theory to evaluate mutual fund performance. As indicated by Brandouy et al. (2012), since the pioneering study by Sengupta (1989), who was probably the first to introduce an explicit efficiency measure into a Mean-Variance (MV) portfolio model, an increasing number of contributions in this particular field have been found in the literature, seeking to provide an alternative to *traditional* mutual fund performance.

Nonparametric frontier alternatives such as Data Envelopment Analysis (DEA; see Charnes et al., 1978) or its nonconvex variant (Free Disposal Hull, FDH; see Deprins et al., 1984) enable an extension of the traditional mean-variance framework to incorporate further dimensions of evaluation (rather than mean and variance alone). When evaluating a portfolio, the aspects that investors want to minimize (such as risk) will be considered as inputs and those to be maximized (such as return), outputs. Factoring in this information, these methods yield the so-called (*efficiency*) scores which summarize the performance of the fund. This is a highly attractive property, especially when considering that alternative investment returns often have skewed distributions so that mean and variance (and possibly the performance indexes relying on these two moments) will be insufficient to properly evaluate the performance of mutual funds.

As a result of these advantages, since the seminal paper of Murthi et al. (1997) was published, the number of contributions in this particular field has grown remarkably, some of which have recently been reviewed by Brandouy et al. (2012), Glawischnig and Sommersguter-Reichmann (2010) and Kerstens et al. (2011a). According to Brandouy et al. (2012), this growing literature can be classified in the following categories: (i) models directly transposed from production theory; (ii) models combining traditional performance measures such as those referred to above with additional dimensions; (iii) models directly transposed from portfolio theory; (iv) hedonic price models. In the particular case of ethical funds, the number of contributions is much lower, basically including Basso and Funari (2003). In the case of Islamic funds, as far as we know there is no empirical evidence of these methods yet, although in the case of Islamic banking several research papers have considered these methods.

As indicated by Glawischnig and Sommersguter-Reichmann (2010), most studies in this field have considered nonparametric methods such as DEA and FDH, their non-convex counterpart.

⁵Among which we would also find Omega, the Sortino ratio, Kappa 3, the upside potential ratio, the Calmar ratio and the modified Sharpe ratio.

However, the most recent contributions have considered more state-of-the-art methods such as the partial frontier approaches order- m (Cazals et al., 2002) and order- α (Aragon et al., 2005) also including partial frontier methods. See, for instance, Daraio and Simar (2006b) or, more recently, Matallín-Sáez et al. (2014).

3.1. Estimating mutual fund performance: FDH and its robust variants

In the first stage of our analysis we will use two efficiency models based on Free Disposal Hull (FDH) (which was initially proposed by Deprins et al., 1984) which, as indicated in the introduction are known in the literature as order- m (Cazals et al., 2002) and order- α (Aragon et al., 2005). In the particular setting of the performance of mutual funds they have already been considered, both in more theoretical contributions, such as those by Daraio and Simar (2006a) and Daraio and Simar (2007b), as well as in recent applications, such as Matallín-Sáez et al. (2014). In these and related contributions the authors provide the reader with some of the benefits of considering not only FDH but also its two robust variants—order- m and order- α .

These types of frontiers—order- m and order- α —offer several advantages over DEA and FDH. Specifically, DEA and FDH are highly sensitive to extreme values and noise in the data, whereas order- m and order- α are not. In addition, they do not impose the convexity assumption (as is the case with DEA) and they have several desirable properties that make them useful for drawing inferences about efficiency. The asymptotic properties of both DEA and FDH also show that they have slow rates of convergence, reflecting the curse of dimensionality (see Simar and Wilson, 2008, p.441), which is usually the case among nonparametric estimators. In addition, order- m and order- α methods allow corrections for the impact of outlying observations which in the current context might arise from atypical observations.

We will only present an intuitive description of the model(s). More formal discussions can be found in, for instance, Cazals et al. (2002), Daraio and Simar (2007b) and Simar and Wilson (2008), whereas a graphical representation is provided by Müller (2008).

3.1.1. The Free Disposal Hull model

The seminal paper by Farrell (1957) inspired researchers to consider performance as a relative estimate, whereby any given decision-making unit (DMU) is evaluated against a frontier of best practice observations. Consider n DMUs (in our case, mutual funds), which are using p heterogeneous and non-negative inputs \mathbf{x} ($x_1 \dots, x_p$) to produce q heterogeneous and non-negative outputs \mathbf{y} (y_1, \dots, y_q). The FDH model relies on a free disposability assumption of inputs and outputs: $\forall(\mathbf{x}, \mathbf{y}) \in \Psi$, if $\tilde{\mathbf{x}} \geq \mathbf{x}$ and $\tilde{\mathbf{y}} \leq \mathbf{y}$ then $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \in \Psi$ (where Ψ denotes the production technology set: $\Psi = \{(\mathbf{x}, \mathbf{y}) | \mathbf{x} \in \mathbb{R}_+^p, \mathbf{y} \in \mathbb{R}_+^q, (\mathbf{x}, \mathbf{y}) \text{ is feasible}\}$).⁶ The best practice production set is defined as a free disposable hull of undominated input-output combinations:⁷

$$\hat{\Psi} = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^{p+q} | (\mathbf{x}, \mathbf{y}) \text{ are attainable}\} \quad (1)$$

The output-oriented efficiencies for the evaluated fund $\hat{\theta}(\mathbf{x}_0, \mathbf{y}_0)$, which maximizes outputs for given inputs, are obtained by estimating the distance to the best practice frontier:

$$\hat{\theta}(\mathbf{x}_0, \mathbf{y}_0) = \sup\{\theta | (\mathbf{x}_0, \theta \mathbf{y}_0) \in \hat{\Psi}_{FDH}\}. \quad (2)$$

An efficient fund obtains an efficiency $\hat{\theta} = 1$, while an inefficient fund yields $\hat{\theta} > 1$. Given the

⁶In other words, if a particular input-output combination (\mathbf{x}, \mathbf{y}) is feasible, it should also be possible to produce \mathbf{y} with more inputs and to produce less outputs with a given input set \mathbf{x} .

⁷Please note that we will refer to the multivariate case (multiple inputs/multiple outputs) by using vector notation throughout the article.

output orientation adopted, $(1 - \hat{\theta})$ should be interpreted as the potential percentage increase in output for the evaluated fund to catch up with the best (efficient) funds.

3.1.2. The robust FDH models

According to FDH, the efficiency measure is obtained by comparison with the full frontier of all observations, defining the maximum output that is technically feasible with a given level of inputs. However, as can be observed from Equation (1), the FDH frontier in the standard FDH setting of Deprins et al. (1984) is deterministic in the sense that all observations from the sample χ potentially constitute the frontier $\hat{\Psi}_{FDH}$.

Alternatively, according to the order- m estimators, what will actually be used as a benchmark is the *expected* maximum output achieved by any m funds chosen randomly from the population, which employs, at most, input level \mathbf{x}_0 . This was initially suggested by Cazals et al. (2002) and allows for mitigating the impact of outlying observations by drawing with replacement subsamples of size $m < n$ among those funds with fewer inputs than the evaluated fund (i.e., among those \mathbf{y}_i so that $\mathbf{x}_0 \geq \mathbf{x}_i$). Performance is then assessed relative to this smaller sample. Therefore, for any \mathbf{y} , the *expected* maximum level will be defined as:

$$\mathbf{y}^\partial = \tilde{\theta}\mathbf{y}. \quad (3)$$

When we choose a high value for m ($m \rightarrow \infty$), the order- m estimator gives the same benchmark as FDH, yielding the same results and, therefore, the most interesting cases will be those for which m is finite. In these cases the order- m does not envelop *all* the data, being more robust to the likely presence of outliers. These outliers which, in the particular output-oriented case we are dealing with will have an efficiency score lower than 1, should be considered as super-efficient with respect to the order- α frontier level.

In contrast to either FDH or DEA, the efficiency scores yielded by order- m are not bounded by 1. In these cases, values equal to unity correspond to *efficient* funds (i.e. those with the best performance), whereas values higher than unity correspond to inefficient funds. Under order- m one may find values for θ lower than one, signalling that the fund operating at the level (\mathbf{x}, \mathbf{y}) is more efficient than the average of m peers randomly drawn from the population of funds with fewer inputs than \mathbf{x} . Formally, the proposed algorithm (Cazals et al., 2002) to compute the order- m estimator has the following steps, for n funds ($i = 1, \dots, n$):⁸

1. For a given level of \mathbf{x}_0 , draw a random sample of size m with replacement among those \mathbf{y}_i , such that $\mathbf{x}_i \leq \mathbf{x}_0$.
2. Obtain the efficiency measures, $\tilde{\theta}_i$.
3. Repeat steps 1 and 2 B times and obtain B efficiency coefficients $\tilde{\theta}_i^b$ ($b = 1, 2, \dots, B$). The quality of the approximation can be tuned by increasing B , but in most applications $B = 200$ seems to be a reasonable choice.
4. Compute the empirical mean of B samples as:

$$\bar{\theta}_i^m = \frac{1}{B} \sum_{b=1}^B \tilde{\theta}_i^b \quad (4)$$

⁸This algorithm illustrates the main difference between FDH and order- m estimators. In the case of order- m , the observation initially selected, $(\mathbf{x}_0, \mathbf{y}_0)$ is compared *only* to the m funds randomly drawn from the population of funds using less inputs than the unit being evaluated. Therefore, it consists of an *expectation* of the maximum achievable output instead of the absolute maximum achievable output that we would consider under FDH.

The order- α quantile-type frontiers share some of the underpinnings of order- m . In the case of order- α the frontier is determined by first fixing the probability $(1 - \alpha)$ of observing points above the order- α frontier. Therefore, under order- α we choose the proportion of funds lying directly below the frontier.

The order- α partial frontiers, originally proposed by Aragon et al. (2005) for the univariate case, were extended to the multivariate case by Daouia and Simar (2007) and, likewise for order- m type frontiers, they have better properties than either DEA or FDH.⁹ However, the main advantage of order- α estimators is the same as that of order- m , i.e. the fact that in finite samples, order- α estimators do not envelop all the data and they are therefore more robust to outliers than FDH and DEA.

In the case of order- α quantile frontiers the benchmark is the output level not exceeded by $(1 - \alpha) \times 100\%$ of funds among the population of funds providing input levels of at least \mathbf{x} .¹⁰

Following Simar and Wilson (2008), for $\alpha \in (0, 1]$, the α -quantile output efficiency score for the mutual fund operating at $(\mathbf{x}, \mathbf{y}) \in \Psi$ can be defined as

$$\theta_\alpha(\mathbf{x}, \mathbf{y}) = \sup\{\theta | F_{\mathbf{y}|\mathbf{x}}(\theta \mathbf{y} | \mathbf{x}) > 1 - \alpha\} \quad (5)$$

We have that $\theta_\alpha(\mathbf{x}, \mathbf{y})$ converges to the FDH estimator $\theta(\mathbf{x}, \mathbf{y})$ when $\alpha \rightarrow 1$. In cases where $\theta_\alpha(\mathbf{x}, \mathbf{y}) = 1$, the fund is “efficient” at the level $\alpha \times 100\%$, since it is dominated by mutual funds providing less input than \mathbf{x} with probability $1 - \alpha$ (Daraio and Simar, 2007a). In those cases where $\theta_\alpha(\mathbf{x}, \mathbf{y}) > 1$ then the unit (\mathbf{x}, \mathbf{y}) has to increase its output to the level $\theta_\alpha(\mathbf{y}, \mathbf{y})\mathbf{x}$ to achieve the output efficient frontier of level $\alpha \times 100\%$. We can also apply the plug-in principle to obtain an intuitive nonparametric estimator of $\theta_\alpha(\mathbf{x}, \mathbf{y}) = 1$ by replacing $F_{\mathbf{y}|\mathbf{x}}(\cdot | \cdot)$ with its empirical counterpart to obtain:

$$\hat{\theta}_{\alpha,n}(\mathbf{x}, \mathbf{y}) = \sup\{\theta | \hat{F}_{\mathbf{y}|\mathbf{x},n}(\theta \mathbf{y} | \mathbf{x}) > 1 - \alpha\} \quad (6)$$

3.2. Analyzing mutual fund performance persistence using partial frontiers

Given the broad universe of available funds, an investor is interested in having tools to select those with better future performance. One possible procedure is to analyse whether funds with worse (better) performance in the past remain so in the future. This information could be used by investors as a guide to plan their investments. In short, it is interesting to analyse the persistence of funds’ performance.

To carry out this analysis we apply a methodology similar to that proposed initially by Carhart (1997) and subsequently used in a high number of studies on performance persistence such as Bollen and Busse (2004), Kosowski et al. (2006), Busse et al. (2010) and Fama and French (2010), among others. This methodology evaluates the performance of portfolios that are built following an investment strategy based on the past performance of funds. If there is persistence one would expect that when investing in the worst (best) funds in the past, if they remain so in the future, one would obtain worse (better) future performance.

To apply this methodology, which the literature refers to as a recursive portfolio, we first estimate the performance of funds in each sample year through the appropriate methodology, as

⁹For instance, they are \sqrt{n} -consistent estimators of the full frontier, since the order of the frontier is allowed to grow with sample size, they are asymptotically unbiased and normally distributed with a known expression for the variance (see Aragon et al., 2005) and, in addition, it can be shown (see Daouia and Simar, 2007) that the order- α frontiers are more robust to extremes than the order- m frontiers (see Daraio and Simar, 2007a, p.74).

¹⁰The monograph by Daraio and Simar (2007a) provides an excellent and comprehensive review of the non-parametric frontier methods considered in this paper, both for the univariate and multivariate cases, as well as illustrating input and output orientations. We have attempted to find a reasonable balance between the amount of information provided in the paper and what can be found in relevant references, not only Daraio and Simar (2007a) but also Simar and Wilson (2008), for instance.

described above. Then funds are ranked according to their performance, from lowest to highest. They are then clustered into deciles, the first decile corresponding to those funds with the worst performance (although the efficiency value would be higher, since we are adopting an output orientation) and the tenth decile including the 10% of funds with better performance (corresponding to the lowest efficiency scores). Then we create portfolios that follow an investment strategy consisting of selecting funds according to their decile in the previous year. For example, portfolio D_1 (D_{10}) invests proportionally and in an equally weighted fashion in funds that were included in the first (tenth) decile in the year before, i.e. last year's worst (best) funds.

Each of these portfolios starts investing in the second year based on the efficiency of the first year, in such a way that the portfolio is reassessed at the beginning of each year and until the end of the sample period. Therefore, from the second year onwards and for each portfolio and posterior years, we identify which are the funds in which we invest. With this information, we compute the monthly return of each portfolio and from the returns it is possible to calculate the values for different variables such as the average return, standard deviation, kurtosis and skewness. In the same way, it is possible to compute the expenses corresponding to each of these portfolios. Then, from these values, which correspond to what will be defined as inputs and outputs, it will be possible to estimate the efficiency of these portfolios. Should persistence be present (at least up to a certain extent) then a given portfolio investing in the worst (best) past funds (such as, for instance, D_1 (D_{10})) would obtain a poor (fair) performance.

In addition, it is convenient to provide results with a certain degree of statistical precision, i.e. whether we can claim that the obtained persistence (or its absence) for the portfolios is significant or not. This would entitle us to ascertain whether the performance of a given portfolio is actually attributable to an investment strategy focusing on past performance, or if this performance is obtained by simply investing in a certain group of funds. Therefore, in order to obtain this statistical precision, we simulate portfolios which do not follow any particular investment strategy but rather invest in our sample funds randomly and equally weightily.

For this, we follow a procedure similar to that considered when building portfolios with strategies based on investing according to past performance. However, in this case funds are not ranked according to past performance but randomly. By the end of each year, the simulated portfolio is reassessed, investing and disinvesting in the sample funds without following any particular investment strategy. Accordingly, we build 2,000 simulated portfolios and for each of them we calculate both the monthly returns and expenses, which in turn allows us to calculate the values for those variables that will be described as inputs and outputs. In the final stage, we re-compute the performance of each portfolio following the partial frontier methods presented above.

The performance of these 2,000 simulated portfolios is not due to any investment strategy in particular, but rather to investing randomly in the sample funds. Therefore, their performance would correspond to a probability distribution, which we compare with the performance of the simulated portfolios built according to past performance—i.e. *random investment strategies* vs. *past performance investment strategies*. This comparison entitles us to conclude whether the results obtained following these two different strategies differ *significantly*—although this procedure is not a formal statistical test. Therefore, considering a given significance level (say 5%), it is possible to conclude that persistence exists when the performance obtained following past performance investment strategies lies outside the 95% central interval of the performance distribution of the 2,000 simulated portfolios (random strategies).

4. Data and performance measurement

4.1. Data sources

The database on SRI and Islamic equity funds that we use is Lipper and the sample period covers from 2001 to 2011. The evolution since 1989 is illustrated in Figure 1, where the number of each type of fund is represented in different axes. In both cases (SRI and Islamic funds) a sharp increase can be observed, which has been particularly high during the “golden decade” which started in 1995 and intensified during the first half of the 2000s. However, during this decade many of the funds that were born did not survive. In order to avoid the bias generated by this phenomenon (i.e. in order to avoid the survivorship bias), our sample is made of all funds and, therefore, it includes new, survivors and non-survivors funds. During the crisis years, which started in 2007, there has been a substantial decline, which has equally affected both types of fund. This type of strategy (i.e. including all data), avoids some pernicious effects such as that of including funds with a relatively low number of observations (which affects the robustness of the performance results), but it comes at the price of comparing funds with different living periods, which could result in some noise in the estimates.

Given these conflicting views, we selected a period and sample of mutual funds with the aim of reaching a balance between the effects referred to above. Consequently, the sample period is confined to the period between December 2000 and March 2011, requiring that the data for any fund in the sample will cover at least 50% of the 123 months of that period. Following these criteria, 9.8% of Islamic funds have to be removed and, therefore, the initial number of 153 Islamic funds is reduced to 138. For SRI funds, 22% of the funds are removed, dropping from 815 to 636. For the final sample, monthly returns were calculated as the variation relative to the monthly net asset value (NAV). Table 1 reports some basic statistics about the mutual funds sample.

Considering Lipper data, mutual funds are grouped by the *geographical area* of their investments. East refers to the Pacific and Asia-Pacific area, Middle refers to Middle East, North Africa and GCC countries,¹¹ West refers to Europe and America and global does not have any particular area of reference. We do this because the funds’ return is the result of both *passive* and *active* management: whereas the return from the latter is the value added by managers above the return from passive management, the return from the former hinges critically on the fund’s investment objectives.

The descriptive statistics on mutual funds for the different areas is also reported in Table 1. As it is apparent in the table, there are no SRI funds which have the Middle East as their investment target area. Comparing SRI and Islamic funds, the averages corresponding to the percentage of data over the sample period and age are relatively similar. However, regardless of the area-objectives considered, the average expenses of Islamic funds are higher than those corresponding to SRI funds. In the case of the annual gross return, for those funds focusing on the West the results are similar, but for others there are remarkable discrepancies. Of special note is the low annual gross return for Islamic funds focusing on the Middle region (1%).

Since the expenses of SRI funds are less than those of the Islamic, the difference between them is somewhat greater if comparing net returns obtained from NAVs. For net returns, Table 1 reports not only the mean but also the median of the cross-sectional distribution of the means of the returns of any fund in each area-objective. In some particular instances the differences are driven by mutual funds on the tails of the distribution of returns. However, discrepancies are remarkable, for some regions in particular. Given the risk component also present in our data, Table 1 shows the average and median of the distribution of the standard deviation for the returns of the funds as well. In this case, the risk levels are similar for both types of funds.

¹¹Gulf Cooperation Council countries, which include Bahrain, Kuwait, Oman, Qatar, Saudi Arabia and the United Arab Emirates.

Finally the last two columns report the average of the skewness and kurtosis of the distributions of the monthly net returns of the funds in each area-objective.

In the specific case of the returns for all funds, the medians of the distribution of the means of net returns for Islamic and SRI funds are 4.48% and 7.18%, respectively. In contrast, the median of the distribution of the standard deviation is *slightly* higher for SRI funds, 20.57%, with respect to 20.14% for Islamic. For the 2.70% of difference, 0.45% is due to expenses (1.53% minus 1.08%) and the rest, 2.25%, is largely due to active management and the differing levels of risk, or area-objective investment, can be attributed it to passive management. However, although the prior descriptive analysis drives us to approximate the behaviour of the funds, it is necessary to conduct a more rigorous analysis of fund performance—and, in this particular study, performance persistence, which will be the objective of the following sections.

4.2. Input and output selection

As indicated in previous sections, one of the main advantages of using frontier techniques to evaluate the performance of mutual funds is their ability to handle several dimensions of mutual fund performance, which are modeled as multiple inputs and outputs. This represents an interesting advantage since it is then possible to include several inputs and thus consider different risk measures apart from standard deviation, or other outputs apart from the traditional mean return measure (such as expected return or the expected excess return).

However, the appropriate choice of dimensions to be modelled, which in the case of frontier techniques corresponds to the selection of inputs and outputs, is crucial. Nguyen-Thi-Thanh (2006) argues that while some investors might be more concerned with central tendencies (mean, standard deviation), others may care more about extreme values (skewness, kurtosis). In our particular application, based on previous literature¹² and data availability, as a main output (y_1) we consider the daily mean return over the sample period. Other outputs, such as skewness (y_2), have also been computed from the daily returns distribution. As inputs, the risk of the fund is measured by the standard deviation of the daily returns (x_1), as well as kurtosis (x_2), also computed from the daily returns. In some of the proposed models the management costs of the fund are also considered as an input. In order to include this type of cost, we consider two variables. First are the fees paid from the fund to managers; second are the loads, including fees and other costs incurred for operational management, e.g. for turnover. Both variables are measured as percentages (average of the sample period) of costs over the managed portfolio size and included in the variable “expenses” (x_3). Descriptive statistics for inputs and outputs are

¹²In their seminal paper, for computing their portfolio efficiency index, Murthi et al. (1997) considered the standard deviation of returns, expense ratio, loads and turnover as inputs and mean gross return as output. Choi and Murthi (2001) made the same choice, although adopting a different DEA formulation, whereas Wilkens and Zhu (2001) considered standard deviation and percentage of periods with negative returns as inputs and mean return, minimum return and skewness as outputs. In Joro and Na (2002) there is an extension of the traditional mean-variance framework using DEA and their methodology includes the risk and cost associated with the transaction as inputs and return and skewness are included as outputs. Chang (2004) proposed a new non-standard DEA formulation based on minimum convex input requirement set: the standard deviation, β , total assets and loads, while the output was the traditional mean return, whereas Brieu et al. (2004) developed a quadratic-constrained (mean-variance) DEA model applying the mean-variance approach with variance as input and mean return as output. More recently, the degree of complexity of some studies has increased, but the selections of inputs and outputs are similar. For instance, Joro and Na (2006) suggested a cubic-constrained mean-variance-skewness framework similar to Brieu et al. (2007), who consider both skewness and mean return as outputs, whereas Lozano and Gutiérrez (2008) proposed a quadratic-constrained DEA model consistent with second-order stochastic dominance (SSD) in order to get an optimal portfolio benchmark for any rational risk-averse investor and Brieu and Kerstens (2009) have introduced a quadratic program that extends the multi-horizon analysis by Morey and Morey (1999) in several ways. Although the literature is now growing, our choice of dimensions to be measured stands with most of the previous applications of these techniques as well as more classic studies on funds’ performance.

reported in table 2.

5. Results

5.1. Expected order- m and order- α efficiency estimates

Tables 3 and 4 provide results on order- m and order- α efficiency estimates for the different mutual funds in our sample. We report some additional statistics apart from the usual ones (not only mean and standard deviation but also the quantiles 25th, 50th and 75th) so that it is possible to obtain enhanced insights on the shape of efficiency scores' distributions. Both Table 3 and Table 4 report the information for several choices of the tuning parameter— m in the case of order- m and α in the case of order- α . The results for each tuning parameter are displayed in different panels in each table. Both Tables 3 and 4 also provide results on performance according to each mutual fund class (SRI or Islamic) as well as its geographical focus—i.e. East, West, Middle and Global—as described in Section 4.

In the case of the order- m class of estimators, we choose different tuning parameters ($m = 25$, $m = 75$ and $m = 150$), since it enables us to ascertain how outliers might affect results; in the extreme case, when $m \rightarrow +\infty$, the class order- m partial frontier yields the same results as FDH. In the case of order- α , a reasonable choice, recommended by Daraio and Simar (2007a), could consist of choosing $\alpha = .90$, $\alpha = .95$ and $\alpha = .99$, similar to the choice of significance levels when running an OLS regression. In this particular paper, though, we have taken advantage of recent developments in the field of efficiency and productivity analysis, which analyse how order- m and order- α indicators are related. Specifically, Daouia and Gijbels (2011) have established that the two classes of indicators are closely related when $\alpha = \alpha(m) = (1/2)^{1/m}$. Therefore, our choice of α parameters is 0.9727, 0.9908 and 0.9954, which correspond to $m = 25$, $m = 75$ and $m = 150$, respectively.

One of the advantages of considering two classes of partial frontiers along with different tuning parameters is that it is possible to check out the robustness of results. The results indicate that, *on average*, SRI funds perform better than Islamic funds. This result is robust across tuning parameters (i.e. different values of m and α) and across partial frontier classes (i.e. for both order- m and order- α). Recall that higher efficiency values represent worse performance. Also, as one could *a priori* expect, because of the characteristics of both order- m and order- α , the higher the tuning parameters, the higher the average inefficiency. Specifically, for $m = 25$, $m = 75$ and $m = 150$ the average efficiency obtained for all funds is 1.0865, 1.1926 and 1.2340, respectively; in the case of order- α , for $\alpha = .9727$, $\alpha = .9908$ and $\alpha = .9954$ (which, following Daouia and Gijbels (2011), would be the equivalent to $m = 25$, $m = 75$ and $m = 150$) these values would be 1.1340, 1.2378 and 1.2802, respectively.

Results are not entirely coincidental when examining different parts of the distribution. In this case, for the worst funds (which would correspond to the 3rd quartile, or 75th quantile) the SRI funds still outperform the Islamic ones. However, although the result is, in general, robust across tuning parameters and partial frontier classes, in the case of order- α the differences shrink substantially. In contrast, in the case of the best funds (which would correspond to the 1st quartile, or 25th quantile), the differences are negligible, although this is partly due to the high number of efficient funds in the vicinity of 1.

One might therefore think that there is something yet to be explained, since the average performance differences between SRI and Islamic funds, regardless of the tuning parameter and partial frontier class, are much higher than those corresponding to either the best funds (1st quartile, which show little differences between both types of funds) or the worst ones (3rd quartile, which show bigger differences). The explanation could lie in a relatively high number of Islamic funds (compared with SRI funds) which perform poorly (i.e. those which lie beyond the 3rd

quartile of the distribution). Actually, the heterogeneity in the performance of Islamic funds, regardless of the tuning parameter and the partial frontier class, is much higher than that within the category of SRI funds, as shown by very disparate standard deviation values (see last columns of Table 3 and 4), which are at least twice as large in the case of Islamic funds than in the case of SRI funds.

Both Tables 3 and 4 also split the results according to the geographical focus of the fund. Although results can be explored from several angles, in this case we find two parallelisms with the comparative analysis performed for Islamic and ethical mutual funds. Specifically, those funds investing in the Middle East perform, *on average*, worse than the rest. This result is robust across tuning parameters and partial frontier classes, without exception. However, analogously to what we reported in the previous paragraphs, this might be partly explained by a relatively high number of funds focusing on this region whose performance is poor which, besides, are responsible for the high performance heterogeneity within this category. Actually, the standard deviation for this specific geographical focus more than doubles that found for the rest. Therefore, we should emphasize the fact that these results are strongly driven by some funds whose performance is particularly weak, i.e. the fact that some of the worst funds are either Islamic or focus on the Middle East should not mislead us to conclude that the funds in these categories perform, in general, worse than those in other categories.

5.2. Results on mutual fund performance persistence

The methodology proposed in Section 3.2 for measuring performance persistence is applied separately to both ethical and Islamic funds and also for measuring the results obtained using order- m and order- α ; as a consequence, the process was repeated four times and three times for each of the tuning parameters (m and α), totalling 12 repetitions.

Table 5 shows the results achieved using order- m , whereas Table 6 provides analogous results for order- α . Each of the two figures contains three panels, one for each tuning parameter value considered. From left to right both tables show the performance and significance of portfolios that invest following an investment strategy based on past performance (past performance investment strategies). Recall that portfolio D_1 (D_{10}) invests in those funds with worse (better) performance in the past and that higher efficiency scores correspond to worse performance (since we are adopting an output orientation). If there were persistence in mutual fund performance, one would expect the worst and the best funds to persist in their relative positions in the future. Therefore, an investment strategy consisting of investing in the worst (best) funds in the past should result in worse (better) performance.

In the case of order- m results, the first row in each panel of Table 5 reports the results for SRI funds, for which performance improves from portfolio D_1 to D_{10} . This constitutes evidence supporting performance persistence, since investing in the worst (best) funds in the past results in worse (better) future performance—recall that the higher the efficiency score, the worse the performance. This result is robust across the three tuning parameters considered (i.e. the three panels). However, the p -values indicate that performance persistence is mostly significant (at either 5% or 1% significance levels) for those portfolios investing in either the worst (deciles D_1 and D_2) and the best (deciles D_9 and D_{10}) funds. For the other portfolios, results are not always entirely coincidental for the three values of m and the strategy of investing in funds whose performance is average (corresponding to the central deciles) does not result in significantly different performance from that yielded by following random investment strategies.

Table 5 also provides results on performance persistence for Islamic funds (second row of each panel). The pattern found is similar to that obtained for SRI funds, although the significance is slightly worse. For the worst funds, corresponding to the D_1 and D_2 deciles, in the case of $m = 150$, we find significance at the 10% level (although it reaches 5% in the case of D_3). For

either $m = 25$ or $m = 75$ significance is lost for D_1 , although it is better (5%) for D_2 . In the opposite tail of the distribution (D_9 and D_{10}), investing in the best performing funds in the past, results in the best future performance and this result is robust for the three tuning parameters. Actually, in the case of $m = 25$ and $m = 75$ significance is also strong (1%) for D_8 . Therefore, the results indicate that in the case of Islamic funds we also find persistence for both the best and the worst funds—especially for the former. This constitutes completely new evidence, since the empirical evidence available so far had not analysed the case of this particular type of fund.

Table 6 reports the order- α counterpart to the results reported in Table 5, for three values of α . In the case of SRI funds, results are basically similar to those yielded by order- m . There are some subtle differences, though, since the significant persistence obtained when investing in those funds with worse past performance always affects a higher number of deciles— D_1 , D_2 , D_3 are always significant at the 1% level and D_4 are significant at the 5% level (for $\alpha = .9727$ and $\alpha = .9954$) and D_5 at the 10% level (although only for $\alpha = .9954$). In the case of the best funds, the portfolios in D_9 and D_{10} show persistence, which implies that investing in the best funds in the past results in better future performance and this result is significant and robust across values of α .¹³

Table 6 also reports results for Islamic funds (second row of each panel). Results are similar to those found for order- m , although we find some subtleties worth mentioning. The performance of those portfolios which invest in the worst funds is generally worse than that obtained by those which invest in the best ones, which should be interpreted as evidence of persistence. However, similar to what we found for order- m in the case of Islamic funds (Table 5), the significance for the worst funds is poor for some values of α (especially $\alpha = .9954$). This implies that the performance of those funds investing in the worst funds in the previous year could be barely distinguished from that obtained when following strategies that invest randomly. In contrast, in the case of the best funds, persistence is limited, since the pattern from decile D_7 to decile D_{10} does not improve remarkably, although persistence is, in general, significant for both D_9 and D_{10} and across α values (although in the case of $\alpha = .9954$ it is only significant at the 10% level).

In sum, there is evidence supporting mutual fund performance persistence and this result is, in general, robust to the technique chosen—order- m or order- α —as well as the tuning parameters within each technique. In the case of Islamic mutual funds, the evidence supporting performance persistence is slightly weaker, especially for order- α , but in the case of the best funds both SRI and Islamic funds show strong persistence.

6. Conclusions

The development of mutual fund industries has motivated a large body of literature attempting to measure mutual fund performance, which is a relevant issue for both investors and managers due to its non-negligible impact on wealth. Although most of this literature has focused on understanding mutual fund performance based on data from the past, future performance is also a question of great concern. Therefore, both individual and institutional investors are particularly interested in ascertaining those funds that will provide the best future results.

The particular variant of the mutual fund performance literature which has been analyzing this and related questions is referred to as the literature on performance *persistence*, which generally deals with the question of whether the relative performance of mutual funds *persists* from one period to another. This is now a well-established body of research where a number of relevant studies have been published, but the variety of results achieved (as to whether we may conclude

¹³However, in the case of D_9 for $\alpha = .9954$ (lower panel) it is not possible to determine the result due to the high number of efficient funds (with an efficiency score of 1) in this category, which makes it impossible to rank them and clustering them into two categories, so that finally all of them are classified into D_{10} .

if persistence actually exists or not) is also high, as documented in several studies.

Two other variants of the literature on mutual fund performance are those dealing with socially responsible investment funds and Islamic funds, which are both growing at remarkable rates—although the number of studies available for the former is still significantly higher than that for the latter. However, the available studies analyzing performance *persistence* for these types of funds are virtually non-existent. Although in the case of SRI funds some authors have dealt indirectly with this issue, in the Islamic case there is no evidence yet.

In this paper we have proposed a recursive investment approach based on recently developed frontier methods for filling this gap in the literature—i.e. analyzing the persistence for both SRI and Islamic funds. Although these techniques do not dominate the finance literature, in the particular case of mutual fund performance measurement they have several advantages, consisting mainly of their ability to model several dimensions of performance, not necessarily restricted to risk and return. These techniques have rarely been considered for analyzing mutual fund performance *persistence* and even less if we extend this particular combination to the analysis of either SRI or Islamic funds.

We consider this mix of methodologies, which had never been done before, is particularly interesting for several reasons. First, the recursive investment approach we are proposing for measuring performance has only rarely been used in the literature. We refine this method by making it possible to ascertain whether the results are statistically significant or not. Second, combining the recursive investment approach with the nonparametric estimators used in the first stage also has some advantages since, given how easily nonparametric frontier estimators accommodate different tuning parameters, conducting the analysis for several of these parameters makes it possible to add robustness to the results obtained.

Our results indicate that performance persistence actually exists for *some* of the funds under analysis. Specifically, both the past year’s winners and losers persisted in their rankings and for both types of funds analyzed—although in the case of Islamic funds the results for the losers was less clear. In addition, although the techniques employed in the paper lie within the general field of nonparametric frontier techniques, they allowed for a variety of different specifications. Our results were robust across most of them.

The results we obtain might be partly explained by the existence of constraints for the types of funds under analysis which require a stronger effort from managers in order to select those investments which, while meeting these constraints, contribute obtaining a better overall performance for the fund. Actually, the previous literature showed that, in the aggregate, there are no significant differences between SRI funds and conventional funds. This would imply that, in general, managers are able to offset the effects of constraints on some particular investments (which might be interesting due to its combination of risk and return) via a stronger effort in selecting assets with good performance.

Therefore, due to the relevant role (in terms of value added) of the managers of constrained funds, one might expect that the effect of their active management on performance will be higher. Indeed, while in the case of conventional mutual funds the evidence of persistence is not conclusive, our results for SRI funds do actually show persistence for both the best and the worst funds. In other words, the increased performance yielded by this extra effort necessary for managing funds with constraints is consistent and is repeated in the future—i.e. it *persists*. However, in the case of the worst funds managers face more difficulties to add value to overcome the effects of constraints and improve performance and this outcome also persists over time.

In the particular case of the Islamic funds, the evidence for the best funds is analogous to that found for the SRI funds, i.e. their managers are interested in adding value, which also persists over time. However, in the case of the worst funds this persistence is not mimicked, which might be indicating that the relatively worse performance might not be driven by their

active management but by other external factors such as the evolution of markets. In this respect, recall that, opposite to the case of SRI funds (among which there are no funds investing in the Middle East), in the case of Islamic funds there is a remarkable number of funds investing in this region and these are precisely those with the worst relative performance.

Therefore, we consider our study has implications in several dimensions. From a methodological point of view, it combines different fields of research which, as far as we know, had barely been combined—i.e. the analysis of mutual fund performance persistence and efficiency estimation using nonparametric robust techniques. The results we obtain are relevant for other dimensions as well (not only the academic community but also practitioners, investors and the market), since we focus on particular types of funds (SRI and Islamic funds) whose performance *persistence* had barely been studied. Our results show that, although last year’s performance only persists for the best and worst funds, the result are significant and, in addition, robust across methodologies (which was only possible due to the techniques used). This persistence is particularly important in financial crises such as the one we are still in that broke in August 2007, during which many investors might be more interested in funds whose performance is more *stable*. Although both SRI and Islamic funds had always been thought of as being more stable than conventional funds, our paper provides compelling evidence that this is actually the case only for the best and worst funds. Financial stability would consequently be closer to what these types of funds offer compared to conventional funds but, given the importance of this topic, some more research in the field would be welcome in order to either reinforce or refute our findings. In-depth analysis of the persistence of poor performers, coinciding with some of Carhart’s (1997) results, as he found that “only the strong, persistent underperformance by the worst-return mutual funds remains anomalous”.

References

- Ahmad, Z. and Ibrahim, H. (2002). A study of performance of the KLSE Shariah Index. *Malaysian Management Journal*, 6:25–34.
- Albaity, M. and Ahmad, R. (2008). Performance of Syariah and composite indices: Evidence from Bursa Malaysia. *Asian Academy of Management Journal of Accounting and Finance*, 4(1):23–43.
- Aragon, Y., Daouia, A., and Thomas-Agnan, C. (2005). Nonparametric frontier estimation: A conditional quantile-based approach. *Econometric Theory*, 21(2):358–389.
- Askari, H., Iqbal, Z., Krichene, N., and Mirakhor, A. (2010). *The Stability of Islamic Finance*. Wiley Finance. Wiley, Singapore.
- Basso, A. and Funari, S. (2003). Measuring the performance of ethical mutual funds: a DEA approach. *Journal of the Operational Research Society*, 54(5):521–531.
- Bauer, R., J. Derwall, J., and Otten, R. (2007). The ethical mutual fund performance debate: New evidence from Canada. *Journal of Business Ethics*, 70:111–124.
- Bauer, R., Koedijk, K., and Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7):1751–1767.
- Binmahfouz, S. (2012). *Investment Characteristics of Islamic Investment Portfolios: Evidence from Saudi Mutual Funds and Global indices*. PhD thesis, Durham University.
- BinMahfouz, S. and Hassan, M. K. (2012). A comparative study between the investment characteristics of Islamic and conventional equity mutual funds in Saudi Arabia. *The Journal of Investing*, 21(4):128–143.
- Bollen, N. P. B. and Busse, J. A. (2004). Short-term persistence in mutual fund performance. *Review of Financial Studies*, 18(2):569–597.
- Brandouy, O., Kerstens, K., and Van de Woestyne, I. (2012). Backtesting super-fund portfolio strategies founded on frontier-based mutual fund ratings. In Pasiouras, F., editor, *Efficiency and Productivity Growth in the Financial Services Industry*. Wiley, New York.
- Briec, W. and Kerstens, K. (2009). Multi-horizon Markowitz portfolio performance appraisals: a general approach. *OMEGA*, 37(1):50–62.
- Briec, W., Kerstens, K., and Jokung, O. (2007). Mean-variance-skewness portfolio performance gauging: A general shortage function and dual approach. *Management Science*, 53(1):135–149.
- Briec, W., Kerstens, K., and Lesourd, J. B. (2004). Nonparametric tests of portfolio investment efficiency: A shortage function generalization. *Journal of Optimization Theory and Applications*, 120(1):1–27.
- Brown, S. J. and Goetzmann, W. N. (1995). Performance persistence. *Journal of Finance*, pages 679–698.
- Busse, J. A., Goyal, A., and Wahal, S. (2010). Performance and persistence in institutional investment management. *Journal of Finance*, 65(2):765–790.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82.

- Cazals, C., Florens, J.-P., and Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106:1–25.
- Chang, K. P. (2004). Evaluating mutual fund performance: an application of minimum convex input requirement set approach. *Computers and Operations Research*, 31(6):929–940.
- Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6):429–444.
- Chen, H., Jegadeesh, N., and Wermers, R. (2000). The value of active mutual fund management: An examination of the stockholdings and trades of fund managers. *Journal of Financial and Quantitative Analysis*, 35:343–368.
- Choi, Y. K. and Murthi, B. P. S. (2001). Relative performance evaluation of mutual funds: A non-parametric approach. *Journal of Business Finance & Accounting*, 28(7 & 8):853–876.
- Cohen, R. B., Coval, J. D., and Pástor, L. (2005). Judging fund managers by the company they keep. *Journal of Finance*, 60(3):1057–1096.
- Cremers, K. J. and Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies*, 22(9):3329–3365.
- Daouia, A. and Gijbels, I. (2011). Robustness and inference in nonparametric partial frontier modeling. *Journal of Econometrics*, 161(2):147–165.
- Daouia, A. and Simar, L. (2007). Nonparametric efficiency analysis: A multivariate conditional quantile approach. *Journal of Econometrics*, 140:375–400.
- Daraio, C. and Simar, L. (2006a). A robust nonparametric approach to evaluate and explain the performance of mutual funds. *European Journal of Operational Research*, 175(1):516–542.
- Daraio, C. and Simar, L. (2006b). A robust nonparametric approach to evaluate and explain the performance of mutual funds. *European Journal of Operational Research*, 175(1):516–542.
- Daraio, C. and Simar, L. (2007a). *Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications*. Studies in Productivity and Efficiency. Springer, New York.
- Daraio, C. and Simar, L. (2007b). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. *Journal of Productivity Analysis*, 28(1):13–32.
- Deprins, D., Simar, L., and Tulkens, H. (1984). Measuring labor-efficiency in post offices. In Marchand, M., Pestieau, P., and Tulkens, H., editors, *The Performance of Public Enterprises: Concepts and Measurement*, chapter 10, pages 243–267. North-Holland, Amsterdam.
- Derigs, U. and Marzban, S. (2008). Review and analysis of current Shariah-compliant equity screening practices. *International Journal of Islamic and Middle Eastern Finance and Management*, 1(4):285–303.
- Dharani, M. and Natarajan, P. (2011). Equanimity of risk and return relationship between Shari’ah index and general index in India. *Journal of Economics and Behavioral Studies*, 2(5):213–222.
- Droms, W. G. (2006). Hot Hands, Cold Hands: Does Past Performance Predict Future Returns? *Journal of Financial Planning*, 19(5):60–69.

- Eling, M. and Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking & Finance*, 31(9):2632–2647.
- Elton, E. J., Gruber, M. J., and Blake, C. R. (1996). The persistence of risk-adjusted mutual fund performance. *Journal of Business*, 69(2):133–157.
- Elton, E. J., Gruber, M. J., Das, S., and Hlavka, M. (1993). Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *Review of Financial Studies*, 6(1):1–22.
- Ernst & Young (2011). Islamic Funds & Investments Report 2011: Achieving Growth in Challenging Times.
- Fama, E. F. and French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *Journal of Finance*, 65(5):1915–1947.
- Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, Ser.A,120:253–281.
- Forte, G. and Miglietta, F. (2007). Islamic mutual funds as faith-based funds in a socially responsible context. Technical report, Social Science Research Network. Available at SSRN: <http://ssrn.com/abstract=1012813> or <http://dx.doi.org/10.2139/ssrn.1012813>.
- Fowler, S. J. and Hope, C. (2007). A critical review of sustainable business indices and their impact. *Journal of Business Ethics*, 76(3):243–252.
- Girard, E. C. and Hassan, M. K. (2008). Is there a cost to faith-based investing: Evidence from FTSE Islamic indices. *The Journal of Investing*, 17(4):112–121.
- Glawischnig, M. and Sommersguter-Reichmann, M. (2010). Assessing the performance of alternative investments using non-parametric efficiency measurement approaches: Is it convincing? *Journal of Banking & Finance*, 34(2):295–303.
- Goetzmann, W. N. and Ibbotson, R. G. (1994). Do winners repeat? *The Journal of Portfolio Management*, 20(2):9–18.
- Gottesman, A. and Morey, M. (2007). Predicting emerging market mutual fund performance. *Journal of Investing*, 16(3):111–122.
- Gregory, A. and Whittaker, J. (2007). Performance and performance persistence of “ethical” unit trusts in the UK. *Journal of Business Finance & Accounting*, 34(7-8):1327–1344.
- Grinblatt, M. and Titman, S. (1992). The persistence of mutual fund performance. *Journal of Finance*, 47(5):1977–1984.
- Hashim, N. (2008). The ftse global islamic and the risk dilemma. Working Paper Series 2008-08, American International University, Office of Research and Publications (ORP), Bangladesh.
- Hayat, R. and Kraeussl, R. (2011). Risk and return characteristics of Islamic equity funds. *Emerging Markets Review*, 12(2):189–203.
- Hendricks, D., Patel, J., and Zeckhauser, R. (1993). Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988. *Journal of Finance*, 48(1):93–130.
- Hiagh, M. and Hazelton, J. (2004). Financial markets: A tool for social responsibility. *Journal of Business Ethics*, 52(1):59–71.

- Hossain, M. Z. (2009). Why is interest prohibited in Islam? A statistical justification. *Humanomics*, 25(4):241–253.
- Hussein, K. and Omran, M. (2005). Ethical investment revisited. *The Journal of Investing*, 14(3):105–126.
- Jensen, M. C. (1969). Risk, the pricing of capital assets, and evaluation of investment portfolios. *Journal of Business*, 42:167–247.
- Joro, T. and Na, P. (2002). Data Envelopment Analysis in mutual fund evaluation: a critical review. Research Report 02-2, Department of Finance and Management Science, School of Business, University of Alberta, Edmonton, Alberta.
- Joro, T. and Na, P. (2006). Portfolio performance evaluation in a mean–variance–skewness framework. *European Journal of Operational Research*, 175(1):446–461.
- Kerstens, K., Mounir, A., and de Woestyne, I. V. (2011a). Non-parametric frontier estimates of mutual fund performance using c- and l-moments: Some specification tests. *Journal of Banking & Finance*, 35(5):1190–1201.
- Kerstens, K., Mounir, A., and Van de Woestyne, I. (2011b). Non-parametric frontier estimates of mutual fund performance using c-and l-moments: Some specification tests. *Journal of Banking & Finance*, 35(5):1190–1201.
- Kosowski, R., Timmermann, A., Wermers, R., and White, H. (2006). Can mutual fund “stars” really pick stocks? New evidence from a bootstrap analysis. *Journal of Finance*, 61(6):2551–2595.
- Lozano, S. and Gutiérrez, E. (2008). Data Envelopment Analysis of mutual funds based on second-order stochastic dominance. *European Journal of Operational Research*, 189(1):230–244.
- Lynch, A. W. and Musto, D. K. (2003). How investors interpret past fund returns. *Journal of Finance*, 58(5):2033–2058.
- Mansor, F. and Bhatti, I. (2011). Risk and return analysis on performance of the Islamic mutual funds: evidence from Malaysia. *Global Economy and Finance Journal*, 4(1):19–31.
- Massa, M. and Patgiri, R. (2009). Incentives and mutual fund performance: Higher performance or just higher risk taking? *Review of Financial Studies*, 22:1777–1815.
- Matallín-Sáez, J. C., Soler-Domínguez, A., and Tortosa-Ausina, E. (2014). On the informativeness of persistence for evaluating mutual fund performance using partial frontiers. *Omega*, 42:47–64.
- Morey, M. R. and Morey, R. C. (1999). Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking. *OMEGA*, 27(2):241–258.
- Müller, J. (2008). Efficiency vectors, efficiency and performance measures: new methods for ranking efficient organizational units. *Journal of Productivity Analysis*, 30(2):99–106.
- Murthi, B. P. S., Choi, Y. K., and Desai, P. (1997). Efficiency of mutual funds and portfolio performance measurement: a nonparametric measurement. *European Journal of Operational Research*, 98:408–418.
- Nguyen-Thi-Thanh, H. (2006). On the use of Data Envelopment Analysis in hedge fund selection. Working paper, Université d’Orléans.

- Novethic (2009). Islamic finance and SRI: any crossover? Unpublished manuscript.
- Quigley, G. and Siquefield, R. (2000). Performance of UK equity unit trusts. *Journal of Asset Management*, 1(1):72–92.
- Renneboog, L., Ter Horst, J., and Zhang, C. (2011). Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation*, 20(4):562–588.
- Sauer, D. A. (1997). The impact of social-responsibility screens on investment performance: Evidence from the Domini 400 Social Index and Domini Equity Mutual Fund. *Review of Financial Economics*, 6(2):137–149.
- Schepers, D. H. (2003). A critique of social investing’s diversity measures. *Business & Society Review*, 108(4):487–508.
- Sengupta, J. K. (1989). Nonparametric tests of efficiency of portfolio investment. *Journal of Economics*, 50(1):1–15.
- Shanmugam, B. and Zahari, Z. (2009). A primer on Islamic finance. Research foundation publications, CFA Institute, Charlottesville.
- Simar, L. and Wilson, P. W. (2008). Statistical inference in nonparametric frontier models: Recent developments and perspectives. In Fried, H., Lovell, C. A. K., and Schmidt, S. S., editors, *The Measurement of Productive Efficiency*, chapter 4, pages 421–521. Oxford University Press, Oxford, 2nd edition.
- Social Investment Forum (2010). *Report on Socially Responsible Investing Trends in the United States*. Available at: <http://ussif.org/resources/research/>.
- Wermers, R. (2003). Is money really ‘smart’? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Available at SSRN: <http://ssrn.com/abstract414420>.
- Wilkens, K. and Zhu, J. (2001). Portfolio evaluation and benchmark selection. *Journal of Alternative Investments*, 4(1):9–19.

Table 1: Descriptive statistics for Islamic and SRI funds distributed by geographical area (2001–2011)

Panel A: Islamic equity mutual funds sample											
Geographical focus ^a	Number	% Date	Age (years)	Annual expenses	Annual gross return	Annual net return		Std.dev. annual net return		Monthly net returns	
						Average	Median	Average	Median	Skewness	Kurtosis
East	51	77%	9.92	1.53%	10.45%	8.92%	9.15%	20.95%	20.35%	−0.47	2.73
Middle	39	68%	7.83	1.69%	1.00%	−0.69%	2.46%	31.59%	17.48%	−1.00	4.89
West	19	82%	10.41	1.41%	7.93%	6.52%	−0.85%	18.11%	28.16%	−0.58	1.86
Global	29	86%	11.22	1.40%	3.70%	2.29%	5.22%	19.19%	18.29%	−1.77	10.75
All funds	138	77%	9.67	1.53%	6.01%	4.48%	4.88%	23.20%	20.14%	−0.91	4.91
Panel B: SRI equity mutual funds sample											
Geographical focus ^a	Number	% Date	Age (years)	Annual expenses	Annual gross return	Annual net return		Std.dev. annual net return		Monthly net returns	
						Average	Median	Average	Median	Skewness	Kurtosis
East	73	75%	10.16	1.26%	13.94%	12.67%	13.20%	24.40%	24.95%	−0.72	2.44
West	350	86%	12.40	0.89%	7.61%	6.71%	5.58%	22.05%	18.84%	−0.86	3.70
Global	213	85%	10.70	1.32%	7.38%	6.06%	4.99%	20.19%	20.83%	−0.87	4.84
All funds	636	84%	11.57	1.08%	8.26%	7.18%	5.84%	21.70%	20.57%	−0.85	3.94

^a East refers to the Pacific and Asia-Pacific area, Middle refers to Middle East, North Africa and GCC countries (Gulf Cooperation Council countries, which include Bahrain, Kuwait, Oman, Qatar, Saudi Arabia and the United Arab Emirates), West refers to Europe and America, and global does not have any particular area of reference.

Table 2: Descriptive statistics for inputs and outputs, Islamic and SRI funds (2001–2011)^a

Class: Islamic		Mean	1 st quartile	Median	3 rd quartile	Std.dev.
Outputs	Gross return (y_1)	0.0051	0.0023	0.0054	0.0093	0.0075
	Skewness (y_2)	-0.9166	-1.0763	-0.5727	-0.2446	1.4745
Inputs	Std.dev. (x_1)	0.0664	0.0499	0.0575	0.0758	0.0280
	Kurtosis (x_2)	5.0176	0.6650	1.5104	3.5814	13.0013
	Expenses (x_3)	0.0153	0.0150	0.0150	0.0175	0.0048
Number of funds: 138						
Class: SRI		Mean	1 st quartile	Median	3 rd quartile	Std.dev.
Outputs	Gross return (y_1)	0.0071	0.0035	0.0061	0.0100	0.0056
	Skewness (y_2)	-0.8614	-1.0260	-0.8229	-0.5883	1.0122
Inputs	Std.dev. (x_1)	0.0620	0.0519	0.0587	0.0687	0.0151
	Kurtosis (x_2)	4.0704	1.6535	2.5617	3.6268	7.9252
	Expenses (x_3)	0.0108	0.0050	0.0113	0.0150	0.0071
Number of funds: 636						

Table 3: Order- m efficiencies, mutual funds (2001–2010)

$m = 25$						
		Mean	1 st quartile	Median	3 rd quartile	Std.dev.
Fund classification	Islamic	1.2030	0.9559	1.0010	1.2536	0.7584
	SRI	1.0643	0.9048	1.0000	1.1939	0.3547
Geographical focus	East	1.2139	0.9864	1.0829	1.3835	0.4475
	West	1.0186	0.8614	0.9805	1.1343	0.3701
	Middle	1.3252	0.9822	1.0000	1.3145	0.9889
	Global	1.1061	0.9413	1.0475	1.2168	0.4272
All funds		1.0865	0.9139	1.0000	1.2004	0.4475
$m = 75$						
		Mean	1 st quartile	Median	3 rd quartile	Std.dev.
Fund classification	Islamic	1.3092	1	1.0724	1.3708	0.8741
	SRI	1.1704	0.9919	1.0686	1.2870	0.3852
Geographical focus	East	1.3076	1.0000	1.1430	1.4955	0.5059
	West	1.1301	0.9807	1.0000	1.2250	0.4057
	Middle	1.4596	1.0000	1.0895	1.4641	1.1628
	Global	1.2060	0.9999	1.1294	1.3073	0.4730
All funds		1.1926	0.9939	1.0692	1.2930	0.4995
$m = 150$						
		Mean	1 st quartile	Median	3 rd quartile	Std.dev.
Fund classification	Islamic	1.3580	1.0000	1.1116	1.4165	0.9296
	SRI	1.2104	1.0000	1.0972	1.3317	0.4109
Geographical focus	East	1.3471	1.0000	1.1747	1.5609	0.5398
	West	1.1680	0.9988	1.0175	1.2693	0.4339
	Middle	1.5237	1.0000	1.1279	1.5168	1.2466
	Global	1.2507	1.0000	1.1654	1.3552	0.4993
All funds		1.2340	1.0000	1.1002	1.3401	0.5320

Notes: the numbers represent descriptive statistics for the efficiency scores yielded by order- m , and for three tuning parameters ($m = 25$, $m = 75$, $m = 150$). Since we are adopting an output orientation, efficiencies closer to one indicate better performance and if a given fund achieves an efficiency score of 1 it implies it is fully efficient (best performance). Efficiency scores lower than 1 indicate that the unit under analysis is an outlier.

Table 4: Order- α efficiencies (maximizing), mutual funds (2001–2010)

$\alpha = .9727$						
		Mean	1 st quartile	Median	3 rd quartile	Std.dev.
Fund classification	Islamic	1.2121	1.0000	1.0000	1.2430	0.7747
	SRI	1.1170	0.9988	1.0393	1.2283	0.3627
Geographical focus	East	1.2308	1.0000	1.0983	1.3882	0.4473
	West	1.0868	0.9802	1.0000	1.1955	0.3962
	Middle	1.3393	1.0000	1.0000	1.3172	1.0167
	Global	1.1265	0.9898	1.0534	1.2283	0.4336
All funds		1.1340	1.0000	1.0302	1.2305	0.4655
$\alpha = .9908$						
		Mean	1 st quartile	Median	3 rd quartile	Std.dev.
Fund classification	Islamic	1.3221	1.0000	1.0869	1.3865	0.8775
	SRI	1.2194	1.0000	1.1235	1.3408	0.4042
Geographical focus	East	1.3284	1.0000	1.1843	1.5217	0.5094
	West	1.1896	1.0000	1.0802	1.2975	0.4441
	Middle	1.4738	1.0000	1.1059	1.4856	1.1588
	Global	1.2305	1.0000	1.1466	1.3333	0.4821
All funds		1.2378	1.0000	1.1190	1.3446	0.5229
$\alpha = .9954$						
		Mean	1 st quartile	Median	3 rd quartile	Std.dev.
Fund classification	Islamic	1.3713	1.0000	1.1246	1.4412	0.9284
	SRI	1.2603	1.0000	1.1604	1.3889	0.4408
Geographical focus	East	1.3678	1.0000	1.2000	1.6027	0.5501
	West	1.2249	1.0000	1.1111	1.3437	0.4791
	Middle	1.5368	1.0000	1.1541	1.5096	1.2481
	Global	1.2808	1.0297	1.1891	1.3841	0.5132
All funds		1.2802	1.0000	1.1558	1.3933	0.5617

Notes: the numbers represent descriptive statistics for the efficiency scores yielded by order- α , and for three tuning parameters ($\alpha = 0.90$, $\alpha = 0.95$, $\alpha = 0.99$). Since we are adopting an output orientation, efficiencies closer to one indicate better performance and if a given fund achieves an efficiency score of 1 it implies it is fully efficient (best performance). Efficiency scores lower than 1 indicate that the unit under analysis is an outlier.

Table 5: Persistence results, order- m

$m = 25$										
	Worst performance					Best performance				
	Decile									
	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
Ethical (p -value)	1.1492 (0.000)	1.1396 (0.000)	1.0197 (0.600)	1.0794 (0.008)	1.0807 (0.007)	1.0195 (0.398)	1.0010 (0.152)	0.9617 (0.005)	0.7547 (0.000)	0.6447 (0.000)
Islamic (p -value)	1.0549 (0.170)	1.0959 (0.041)	1.0842 (0.060)	1.0273 (0.359)	1.0811 (0.070)	1.0203 (0.579)	0.9677 (0.142)	0.7156 (0.002)	0.8767 (0.006)	0.9072 (0.014)
$m = 75$										
	Worst performance					Best performance				
	Decile									
	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
Ethical (p -value)	1.1670 (0.000)	1.1105 (0.001)	1.0890 (0.014)	1.0637 (0.137)	1.0858 (0.020)	1.0280 (0.336)	0.9952 (0.028)	1.0377 (0.492)	0.8566 (0.000)	0.7325 (0.000)
Islamic (p -value)	1.0811 (0.143)	1.1142 (0.048)	1.0999 (0.077)	1.0888 (0.111)	1.0634 (0.249)	0.9173 (0.002)	0.9731 (0.031)	0.9094 (0.002)	0.9584 (0.016)	0.9639 (0.021)
$m = 150$										
	Worst performance					Best performance				
	Decile									
	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
Ethical (p -value)	1.1730 (0.000)	1.1380 (0.000)	1.0661 (0.205)	1.0651 (0.219)	1.0493 (0.436)	1.0384 (0.388)	1.0262 (0.227)	1.0402 (0.417)	1.0000 (0.042)	0.8397 (0.000)
Islamic (p -value)	1.1222 (0.070)	1.1306 (0.057)	1.1476 (0.032)	1.1102 (0.113)	1.0498 (0.415)	1.0882 (0.234)	1.0000 (0.056)	1.0503 (0.419)	0.9528 (0.003)	0.9691 (0.007)

Table 6: Persistence results, order- α

$\alpha = .9727$										
	Worst performance					Best performance				
	Decile									
	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
Ethical (p -value)	1.1490 (0.001)	1.1328 (0.003)	1.1766 (0.000)	1.0654 (0.046)	1.0361 (0.273)	1.0282 (0.590)	1.0884 (0.983)	1.0000 (0.153)	0.9464 (0.001)	0.8924 (0.000)
Islamic (p -value)	1.0438 (0.194)	1.0859 (0.050)	1.0587 (0.115)	1.0697 (0.086)	1.0000 (0.632)	1.0000 (0.405)	1.0089 (0.491)	0.9815 (0.195)	0.8624 (0.003)	0.9264 (0.018)
$\alpha = .9908$										
	Worst performance					Best performance				
	Decile									
	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
Ethical (p -value)	1.1562 (0.020)	1.2384 (0.000)	1.1792 (0.005)	1.0686 (0.148)	1.0349 (0.532)	1.0124 (0.150)	0.9735 (0.001)	1.0518 (0.733)	1.0000 (0.046)	0.9550 (0.000)
Islamic (p -value)	1.1046 (0.070)	1.1057 (0.068)	1.1074 (0.064)	1.0692 (0.227)	1.0684 (0.232)	1.0000 (0.140)	1.0000 (0.140)	– –	1.0598 (0.694)	0.9257 (0.004)
$\alpha = .9954$										
	Worst performance					Best performance				
	Decile									
	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
Ethical (p -value)	1.1595 (0.000)	1.0994 (0.002)	1.0916 (0.005)	1.0716 (0.048)	1.0000 (0.058)	1.0441 (0.358)	0.9697 (0.000)	1.0150 (0.176)	– –	0.9643 (0.000)
Islamic (p -value)	1.1125 (0.101)	1.1048 (0.126)	1.0942 (0.171)	1.1003 (0.146)	1.0470 (0.435)	1.0506 (0.300)	1.0000 (0.075)	1.0000 (0.075)	1.0000 (0.075)	1.0014 (0.080)

Figure 1: Evolution of the number of SRI funds vs. Islamic funds, 1989–2011

